

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/688,069

DATE: 08/09/2001
TIME: 13:27:54

Input Set : A:\SeqList.txt
Output Set: N:\CRF3\08092001\I688069.raw

3 <110> APPLICANT: Subramaniam, S.; Slater, S.; Karberg, K.; Chen, R.; Valentin, H.;
Wong, Y.

5 <120> TITLE OF INVENTION: Nucleic Acid Sequences to Proteins Involved in Tocopherol
Synthesis

7 <130> FILE REFERENCE: 16515.054
9 <140> CURRENT APPLICATION NUMBER: US 09/688,069
10 <141> CURRENT FILING DATE: 2000-10-14
12 <160> NUMBER OF SEQ ID NOS: 114
14 <210> SEQ ID NO: 1
15 <211> LENGTH: 1182
16 <212> TYPE: DNA
17 <213> ORGANISM: Arabidopsis sp.
19 <400> SEQUENCE: 1

ENTERED

2.5

```

21 atggaggtctc tgctctctag ttcttctctt gtttccgctg ctgggtgggtt ttgttggaag 60
22 aagcagaatc taaagctcca ctctttatca gaaatccgag ttctgcgttg tgattcgagt 120
23 aaagtgtgctg caaaaccgaa gtttaggaac aatcttgta ggcctgatgg tcaaggatct 180
24 tcattgttgt tgtatccaaa acataagtcg agatttcggg ttaatgccac tgcgggtcag 240
25 cctgaggctt tcgactcgaa tagcaaacag aagtctttta gagactcgtt agatgcgttt 300
26 tacaggtttt ctaggcctca tacagttatt ggcacagtgc ttagcatttt atctgtatct 360
27 ttcttagcag tagagaaggt ttctgatata tctcctttac ttttcaactgg catcttgag 420
28 gctgtgtgtg cagctctcat gatgaacatt tacatagttg ggctaaatca gttgtctgat 480
29 gttgaaatag ataagggtta caagccctat cttccattgg catcaggaga atattctggt 540
30 aacaccggca ttgcaatagt agcttccctt tccatcatga gtttctggct tgggtggatt 600
31 gttggttcat ggccattggt ctgggctctt tttgtgagtt tcatgctcgg tactgcatac 660
32 tctatcaatt tgccactttt acggtggaaa agatttgcag tgggtgcagc aatgtgtatc 720
33 ctgcgtgtcc gagctattat tgttcaaata gcctttttatc ttagcatttc gacacatgtg 780
34 tttggaagac caatcttggt cactaggcct cttattttcg ccactgcgtt tatgagcttt 840
35 ttctctgtcg ttattgcatt gttaaggat atacctgata tcgaagggga taagatatcc 900
36 ggaatccgat cattctctgt aactctgggt cagaaacggg tgttttgac atgtgttaca 960
37 ctacttcaaa tggcttacgc tgttgcaatt ctagtggag ccacatctcc attcatatgg 1020
38 agcaaagtca tctcgggtgt ggtcatgtt atactcgcaa caactttgtg ggctcgagct 1080
39 aagtcggtg atctgagtag caaaaccgaa ataacttcat gttatatgtt catatggaag 1140
40 ctcttttatg cagagtactt gctgttacct ttttgaagt ga 1182

```

43 <210> SEQ ID NO: 2
44 <211> LENGTH: 393
45 <212> TYPE: PRT
46 <213> ORGANISM: Arabidopsis sp.
48 <400> SEQUENCE: 2

```

50 Met Glu Ser Leu Leu Ser Ser Ser Ser Leu Val Ser Ala Ala Gly Gly
51 1 5 10 15
53 Phe Cys Trp Lys Lys Gln Asn Leu Lys Leu His Ser Leu Ser Glu Ile
54 20 25 30
56 Arg Val Leu Arg Cys Asp Ser Ser Lys Val Val Ala Lys Pro Lys Phe
57 35 40 45
59 Arg Asn Asn Leu Val Arg Pro Asp Gly Gln Gly Ser Ser Leu Leu Leu
60 50 55 60
62 Tyr Pro Lys His Lys Ser Arg Phe Arg Val Asn Ala Thr Ala Gly Gln
63 65 70 75 80

```

65 Pro Glu Ala Phe Asp Ser Asn Ser Lys Gln Lys Ser Phe Arg Asp Ser

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/688,069

DATE: 08/09/2001
TIME: 13:27:54

Input Set : A:\SeqList.txt
Output Set: N:\CRF3\08092001\I688069.raw

```

66                               85                               90                               95
68 Leu Asp Ala Phe Tyr Arg Phe Ser Arg Pro His Thr Val Ile Gly Thr
69                               100                               105                               110
71 Val Leu Ser Ile Leu Ser Val Ser Phe Leu Ala Val Glu Lys Val Ser
72                               115                               120                               125
74 Asp Ile Ser Pro Leu Leu Phe Thr Gly Ile Leu Glu Ala Val Val Ala
76                               130                               135                               140
77 Ala Leu Met Met Asn Ile Tyr Ile Val Gly Leu Asn Gln Leu Ser Asp
79 145                               150                               155                               160
80 Val Glu Ile Asp Lys Val Asn Lys Pro Tyr Leu Pro Leu Ala Ser Gly
82                               165                               170                               175
83 Glu Tyr Ser Val Asn Thr Gly Ile Ala Ile Val Ala Ser Phe Ser Ile
84                               180                               185                               190
86 Met Ser Phe Trp Leu Gly Trp Ile Val Gly Ser Trp Pro Leu Phe Trp
87                               195                               200                               205
89 Ala Leu Phe Val Ser Phe Met Leu Gly Thr Ala Tyr Ser Ile Asn Leu
90                               210                               215                               220
92 Pro Leu Leu Arg Trp Lys Arg Phe Ala Leu Val Ala Ala Met Cys Ile
93 225                               230                               235                               240
95 Leu Ala Val Arg Ala Ile Ile Val Gln Ile Ala Phe Tyr Leu His Ile
96                               245                               250                               255
98 Gln Thr His Val Phe Gly Arg Pro Ile Leu Phe Thr Arg Pro Leu Ile
99                               260                               265                               270
101 Phe Ala Thr Ala Phe Met Ser Phe Phe Ser Val Val Ile Ala Leu Phe
102                               275                               280                               285
104 Lys Asp Ile Pro Asp Ile Glu Gly Asp Lys Ile Phe Gly Ile Arg Ser
105                               290                               295                               300
107 Phe Ser Val Thr Leu Gly Gln Lys Arg Val Phe Trp Thr Cys Val Thr
108 305                               310                               315                               320
110 Leu Leu Gln Met Ala Tyr Ala Val Ala Ile Leu Val Gly Ala Thr Ser
111                               325                               330                               335
113 Pro Phe Ile Trp Ser Lys Val Ile Ser Val Val Gly His Val Ile Leu
114                               340                               345                               350
116 Ala Thr Thr Leu Trp Ala Arg Ala Lys Ser Val Asp Leu Ser Ser Lys
117                               355                               360                               365
119 Thr Glu Ile Thr Ser Cys Tyr Met Phe Ile Trp Lys Leu Phe Tyr Ala
120                               370                               375                               380
122 Glu Tyr Leu Leu Leu Pro Phe Leu Lys
123 385                               390
126 <210> SEQ ID NO: 3
127 <211> LENGTH: 1224
128 <212> TYPE: DNA
129 <213> ORGANISM: Arabidopsis sp.
131 <400> SEQUENCE: 3
133 atggcggtttt ttgggtcttc ccgtgtttca agacggttgt tgaaatcttc cgtctccgta 60
134 actccatctt cttcctctgc tcttttgcaa tcacaacata aatccttgtc caatcctgtg 120
135 actaccatt acacaaatcc ttctactaag tggtatcctt catggaatga taattaccaa 180
136 gtatggagta aaggaagaga attgcatcag gagaagtttt ttggtgttgg ttggaattac 240
137 agattaattt gtggaatgtc gtcgtcttct tcggttttgg agggaaagcc gaagaaagat 300

```

Input Set : A:\SeqList.txt
Output Set: N:\CRF3\08092001\I688069.raw

```

138 gataaggaga agagtgatgg tgttgttgtt aagaaagctt cttggataga tttgtattta 360
139 ccagaagaag ttagagggtta tgctaagctt gctcgattgg ataaacccat tggaacttgg 420
140 ttgcttgctg ggccttgatg gtggtcgatt gcggttgctg ctgacctggg aagccttcca 480
141 agtttttaaat atatggcttt atttggttgc ggagcattac ttcttagagg tgctggttgt 540
142 actataaatg atctgcttga tcaggacata gatacaaagg ttgatcgtag aaaactaaga 600
143 cctatcgcca gtgggtctttt gacaccattt caagggattg gattttctgg gctgcagttg 660
144 ctttttaggct tagggattct tctccaactt aacaattaca gccgtgtttt aggggcttca 720
145 tctttgttac ttgtcttttc ctaccactt atgaagagg ttacattttg gcctcaagcc 780
146 ttttttaggtt tgaccataaa ctggggagca ttgttaggat ggactgcagt taaaggaagc 840
147 atagcaccat ctattgtact cctctcttat ctctcggag tctgctggac ccttgtttat 900
148 gatactattt atgcacatca ggacaaagaa gatgatgtaa aagttgggtg taagtcaaca 960
149 gcccttagat tcggtgataa tacaagctt tggttaactg gatttggcac agcatccata 1020
150 ggttttcttg cactttcttg attcagtgca gatctcgggt ggcaatatta cgcatcactg 1080
151 gccgctgcat caggacagtt aggatggcaa atagggacag ctgacttatc atctggtgct 1140
152 gactgcagta gaaaatttgt gtcgaacaag tggtttggtg ctattatatt tagtggagtt 1200
153 gtacttgga gaagttttca ataa 1224
156 <210> SEQ ID NO: 4
157 <211> LENGTH: 407
158 <212> TYPE: PRT
159 <213> ORGANISM: Arabidopsis sp.
161 <400> SEQUENCE: 4
163 Met Ala Phe Phe Gly Leu Ser Arg Val Ser Arg Arg Leu Leu Lys Ser
164 1 5 10 15
166 Ser Val Ser Val Thr Pro Ser Ser Ser Ser Ala Leu Leu Gln Ser Gln
167 20 25 30
169 His Lys Ser Leu Ser Asn Pro Val Thr Thr His Tyr Thr Asn Pro Phe
170 35 40 45
172 Thr Lys Cys Tyr Pro Ser Trp Asn Asp Asn Tyr Gln Val Trp Ser Lys
173 50 55 60
175 Gly Arg Glu Leu His Gln Glu Lys Phe Phe Gly Val Gly Trp Asn Tyr
176 65 70 75 80
178 Arg Leu Ile Cys Gly Met Ser Ser Ser Ser Ser Val Leu Glu Gly Lys
179 85 90 95
181 Pro Lys Lys Asp Asp Lys Glu Lys Ser Asp Gly Val Val Val Lys Lys
182 100 105 110
184 Ala Ser Trp Ile Asp Leu Tyr Leu Pro Glu Glu Val Arg Gly Tyr Ala
185 115 120 125
187 Lys Leu Ala Arg Leu Asp Lys Pro Ile Gly Thr Trp Leu Leu Ala Trp
188 130 135 140
190 Pro Cys Met Trp Ser Ile Ala Leu Ala Ala Asp Pro Gly Ser Leu Pro
191 145 150 155 160
193 Ser Phe Lys Tyr Met Ala Leu Phe Gly Cys Gly Ala Leu Leu Leu Arg
194 165 170 175
196 Gly Ala Gly Cys Thr Ile Asn Asp Leu Leu Asp Gln Asp Ile Asp Thr
197 180 185 190
199 Lys Val Asp Arg Thr Lys Leu Arg Pro Ile Ala Ser Gly Leu Leu Thr
200 195 200 205
202 Pro Phe Gln Gly Ile Gly Phe Leu Gly Leu Gln Leu Leu Leu Gly Leu
203 210 215 220

```

Input Set : A:\SeqList.txt
Output Set: N:\CRF3\08092001\I688069.raw

```

205 Gly Ile Leu Leu Gln Leu Asn Asn Tyr Ser Arg Val Leu Gly Ala Ser
206 225                230                235                240
208 Ser Leu Leu Leu Val Phe Ser Tyr Pro Leu Met Lys Arg Phe Thr Phe
209                245                250                255
211 Trp Pro Gln Ala Phe Leu Gly Leu Thr Ile Asn Trp Gly Ala Leu Leu
212                260                265                270
214 Gly Trp Thr Ala Val Lys Gly Ser Ile Ala Pro Ser Ile Val Leu Pro
215                275                280                285
217 Leu Tyr Leu Ser Gly Val Cys Trp Thr Leu Val Tyr Asp Thr Ile Tyr
218                290                295                300
220 Ala His Gln Asp Lys Glu Asp Asp Val Lys Val Gly Val Lys Ser Thr
221 305                310                315                320
223 Ala Leu Arg Phe Gly Asp Asn Thr Lys Leu Trp Leu Thr Gly Phe Gly
224                325                330                335
226 Thr Ala Ser Ile Gly Phe Leu Ala Leu Ser Gly Phe Ser Ala Asp Leu
227                340                345                350
229 Gly Trp Gln Tyr Tyr Ala Ser Leu Ala Ala Ala Ser Gly Gln Leu Gly
230                355                360                365
232 Trp Gln Ile Gly Thr Ala Asp Leu Ser Ser Gly Ala Asp Cys Ser Arg
233                370                375                380
235 Lys Phe Val Ser Asn Lys Trp Phe Gly Ala Ile Ile Phe Ser Gly Val
236 385                390                395                400
238 Val Leu Gly Arg Ser Phe Gln
239                405
242 <210> SEQ ID NO: 5
243 <211> LENGTH: 1296
244 <212> TYPE: DNA
245 <213> ORGANISM: Arabidopsis sp.
247 <400> SEQUENCE: 5
249 atgtggcgaa gatctgttgt ttctcgttta tcttcaagaa tctctgtttc ttcttcgtta 60
250 ccaaacccta gactgattcc ttggtcccgc gaattatgtg ccgttaatag cttctcccag 120
251 cctccggtct cgacggaatc aactgctaag ttagggatca ctggtgttag atctgatgcc 180
252 aatcgagttt ttgccactgc tactgccgcc gctacagcta cagctaccac cggtagagatt 240
253 tcgtctagag ttgcgctttt ggctggatta gggcatcact acgctcgttg ttattgggag 300
254 ctttctaaag ctaaacttag tatgcttggt gttgcaactt ctggaactgg gtatattctg 360
255 ggtaögggaa atgctgcaat tagcttcccg gggcttttgt acacatgtgc aggaaccatg 420
256 atgattgctg catctgctaa ttcttgaat cagatttttg agataagcaa tgattctaag 480
257 atgaaaagaa cgatgctaag gccattgcct tcaggacgta ttagtgttcc acacgctgtt 540
258 gcatgggcta ctattgctgg tgcttctggt gcttgtttgt tggccagcaa gactaatatg 600
259 ttggctgctg gacttgcatc tgccaatctt gtactttatg cgtttgttta tactccgttg 660
260 aagcaacttc accctatcaa tacatgggtt ggcgctgttg ttggtgctat cccacccttg 720
261 cttgggtggg cggcagcgtc tggtcagatt tcatacaatt cgatgattct tccagctgct 780
262 ctttactttt ggcagatacc tcattttatg gcccttgca c atctctgccc caatgattat 840
263 gcagctggag gttacaagat gttgtcactc tttgatccgt cagggaagag aatagcagca 900
264 gtggctctaa ggaactgctt ttacatgata cctctcggtt tcatcgcta tgactggggg 960
265 ttaacctcaa gttggttttg cctcgaatca acacttctca cactagcaat cgctgcaaca 1020
266 gcattttcat tctaccgaga ccggaccatg cataaagcaa ggaaaatgtt ccatgccagt 1080
267 cttctcttcc ttcctgtttt catgtctggt cttctcttac accgtgtctc taatgataat 1140
268 cagcaacaac tcgtagaaga agccggatta acaaattctg tatctggtga agtcaaaact 1200

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Input Set : A:\SeqList.txt
Output Set: N:\CRF3\08092001\I688069.raw

```

269 cagaggcgaa agaaacgtgt ggctcaacct ccggtggctt atgcctctgc tgcaccgttt 1260
270 cctttctctcc cagctccttc cttctactct ccatga 1296
273 <210> SEQ ID NO: 6
274 <211> LENGTH: 431
275 <212> TYPE: PRT
276 <213> ORGANISM: Arabidopsis sp.
278 <400> SEQUENCE: 6
280 Met Trp Arg Arg Ser Val Val Tyr Arg Phe Ser Ser Arg Ile Ser Val
281 1 5 10 15
283 Ser Ser Ser Leu Pro Asn Pro Arg Leu Ile Pro Trp Ser Arg Glu Leu
284 20 25 30
286 Cys Ala Val Asn Ser Phe Ser Gln Pro Pro Val Ser Thr Glu Ser Thr
287 35 40 45
289 Ala Lys Leu Gly Ile Thr Gly Val Arg Ser Asp Ala Asn Arg Val Phe
290 50 55 60
292 Ala Thr Ala Thr Ala Ala Thr Ala Thr Ala Thr Thr Gly Glu Ile
293 65 70 75 80
295 Ser Ser Arg Val Ala Ala Leu Ala Gly Leu Gly His His Tyr Ala Arg
296 85 90 95
298 Cys Tyr Trp Glu Leu Ser Lys Ala Lys Leu Ser Met Leu Val Val Ala
299 100 105 110
301 Thr Ser Gly Thr Gly Tyr Ile Leu Gly Thr Gly Asn Ala Ala Ile Ser
302 115 120 125
304 Phe Pro Gly Leu Cys Tyr Thr Cys Ala Gly Thr Met Met Ile Ala Ala
305 130 135 140
307 Ser Ala Asn Ser Leu Asn Gln Ile Phe Glu Ile Ser Asn Asp Ser Lys
308 145 150 155 160
310 Met Lys Arg Thr Met Leu Arg Pro Leu Pro Ser Gly Arg Ile Ser Val
311 165 170 175
313 Pro His Ala Val Ala Trp Ala Thr Ile Ala Gly Ala Ser Gly Ala Cys
314 180 185 190
316 Leu Leu Ala Ser Lys Thr Asn Met Leu Ala Ala Gly Leu Ala Ser Ala
317 195 200 205
319 Asn Leu Val Leu Tyr Ala Phe Val Tyr Thr Pro Leu Lys Gln Leu His
320 210 215 220
322 Pro Ile Asn Thr Trp Val Gly Ala Val Val Gly Ala Ile Pro Pro Leu
323 225 230 235 240
325 Leu Gly Trp Ala Ala Ala Ser Gly Gln Ile Ser Tyr Asn Ser Met Ile
326 245 250 255
328 Leu Pro Ala Ala Leu Tyr Phe Trp Gln Ile Pro His Phe Met Ala Leu
329 260 265 270
331 Ala His Leu Cys Arg Asn Asp Tyr Ala Ala Gly Gly Tyr Lys Met Leu
332 275 280 285
334 Ser Leu Phe Asp Pro Ser Gly Lys Arg Ile Ala Ala Val Ala Leu Arg
335 290 295 300
337 Asn Cys Phe Tyr Met Ile Pro Leu Gly Phe Ile Ala Tyr Asp Trp Gly
338 305 310 315 320
340 Leu Thr Ser Ser Trp Phe Cys Leu Glu Ser Thr Leu Leu Thr Leu Ala
341 325 330 335

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/688,069

DATE: 08/09/2001

TIME: 13:27:55

Input Set : A:\SeqList.txt

Output Set: N:\CRF3\08092001\I688069.raw

L:391 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:392 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:844 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:845 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:846 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:945 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25
L:960 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:979 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27
L:980 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27
L:982 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27
L:2799 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:102